PB82-127499

Computer Technology: Pattern Recognition of Behavioral Events in the Nonhuman Primate

Iowa State Univ.
Ames

Prepared for

Health Effects Research Lab.
Research Triangle Park, NC

1980

U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. RECIPIENT'S ACCESSION NO.
EPA-600/J-80-054	JOURNAL ARTICLE	PB82 127499

4. TITLE AND SUBTITLE	5. REPORT DATE
Computer Technology Pattern Recognition of Behavioral Events in the Nonhuman Primate	September 1981
	6. PERFORMING ORGANIZATION CODE 308088

7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.
W.J. Kernan, Jr., W. J. Higby, D.L. Hopper, W. Cunningham, W.E. Lloyd and L. Reiter	

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT NO.
Veterinary Diagnostic Laboratory Iowa State University Ames, Iowa 50010	C9ZN1D
	11. CONTRACT/GRANT NO. 68-02-2288

12. SPONSORING AGENCY NAME AND ADDRESS	13. TYPE OF REPORT AND PERIOD COVERED
Office of Research and Development (RTP,NC) Health Effects Research Laboratory US Environmental Protection Agency Research Triangle Park, NC 27711	14. SPONSORING AGENCY CODE EPA-600/11

15. SUPPLEMENTARY NOTES

Published In: Behavioral Research Methods & Instrumentation 12(5): 524-534, 1980

16. ABSTRACT

Techniques used in computer graphics and pattern analysis have been applied to the tasks of observing, classifying, and recording spontaneous behavioral activities in the captive primate. The goal in designing this system was to provide a computer-based pattern recognition system capable of identifying normal activity patterns of the primate Macaca fascicularis. If successful, this system can be used to detect alterations in these patterns under conditions of toxicologic and pharmacologic insult. In this article, the development of our system, which is called PROBE (pattern recognition of behavioral events), is discussed.

Throughout this project, the purpose was to solve a unique application problem. Existing pattern recognition methods (Fu, 1968; Fukunaga, 1972; Pavlidis, 1977), consistent with the adequate solution of our problem, were applied.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS (This Report)	21. NO. OF PAGES
RELEASE TO PUBLIC	UNCLASSIFIED	
	20. SECURITY CLASS (This page) UNCLASSIFIED	22. PRICE

EPA-600/J-80-054
Journal Article

PB82-127499

Behavior Research Methods & Instrumentation
1980, Vol. 12 (5), 524-534

COMPUTER TECHNOLOGY

Pattern recognition of behavioral events in the nonhuman primate

W. J. KERNAN, JR., and W. J. HIGBY
Energy and Mineral Resources Research Institute, Iowa State University, Ames, Iowa 50010

D. L. HOPPER, W. CUNNINGHAM, and W. E. LLOYD
Veterinary Diagnostic Laboratory, Iowa State University, Ames, Iowa 50010

and

L. REITER
Neurotoxicology Division, Health Effects Research Laboratory, Environmental Protection Agency
Research Triangle Park, North Carolina 27711

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA 22161

Behavior Research Methods & Instrumentation
1980, Vol. 12 (5), 524-534

COMPUTER TECHNOLOGY

Pattern recognition of behavioral events in the nonhuman primate

W. J. KERNAN, JR., and W. J. HIGBY
Energy and Mineral Resources Research Institute, Iowa State University, Ames, Iowa 50010

D. L. HOPPER, W. CUNNINGHAM, and W. E. LLOYD
Veterinary Diagnostic Laboratory, Iowa State University, Ames, Iowa 50010

and

L. REITER
Neurotoxicology Division, Health Effects Research Laboratory, Environmental Protection Agency Research Triangle Park, North Carolina 27711

A computerized pattern recognition system has been developed that is capable of identifying 40 separate spontaneously occurring behavioral acts of the primate *Macaca fascicularis*. The system, called PROBE (pattern recognition of behavioral events), is described in detail. In its present stage of development, PROBE classifies behavioral activity with a reliability comparable to trained human observers. The potential applications for and improvements to the PROBE system are discussed.

Techniques used in computer graphics and pattern analysis have been applied to the tasks of observing, classifying, and recording spontaneous behavioral activities in the captive primate. The goal in designing this system was to provide a computer-based pattern recognition system capable of identifying normal activity patterns of the primate *Macaca fascicularis*. If successful, this system can be used to detect alterations in these patterns under conditions of toxicologic and pharmacologic insult. In this article, the development of our system, which is called PROBE (pattern recognition of behavioral events), is discussed.

Throughout this project, the purpose was to solve a unique application problem. Existing pattern recognition methods (Fu, 1968; Fukunaga, 1972; Pavlidis, 1977), consistent with the adequate solution of our problem, were applied.

HARDWARE

The computer system consists of a Digital Equipment Corporation PDP-11/34 computer with a floating-point processor and 32K of core memory, a nine-track magnetic tape drive, card reader, line printer, console typewriter, Tektronics Model 4012 display scope, and a 44-million-word disk drive. A schematic view of the PROBE hardware is shown in Figure 1. The card reader and line printer are required for easy program input and analyzed data output, respectively. The magnetic tape drive provides permanent storage for the data. The disk is a vital part of the overall pattern recognition hardware, as the amount of data generated in image processing is enormous, making it essential that a rapid, direct-access, high-volume storage device be used to process the video data. The storage display scope serves as the window by which the human observer can process the same image information as the computer, and as such, it is the only means of checking the reliability of the computer classifications.

With the assistance of graphics displays, the experimenter can interactively participate in the pattern recognition process. Like many other computer-aided systems, this interactive approach allows the speed and accuracy of a computer to be combined with the ability and intuition of the human at various stages of the pattern recognition process.

The cage in which the monkey is observed is viewed by three orthogonal video cameras. The camera signals are processed through a video system (Kantronics) before the data are transferred to the computer. This video system is composed of three (VQ-125) video discriminator units and a video interface (VI-225). Each discriminator accepts a standard composite video input signal (1 V, peak to peak) from a monochrome video camera and converts this signal into binary data representing those portions of the video signal that are within

This research was supported in part by the United States Environmental Protection Agency under E.P.A. Contract 68-02-2288.

0005-7878/80/050524-11$01.35/0

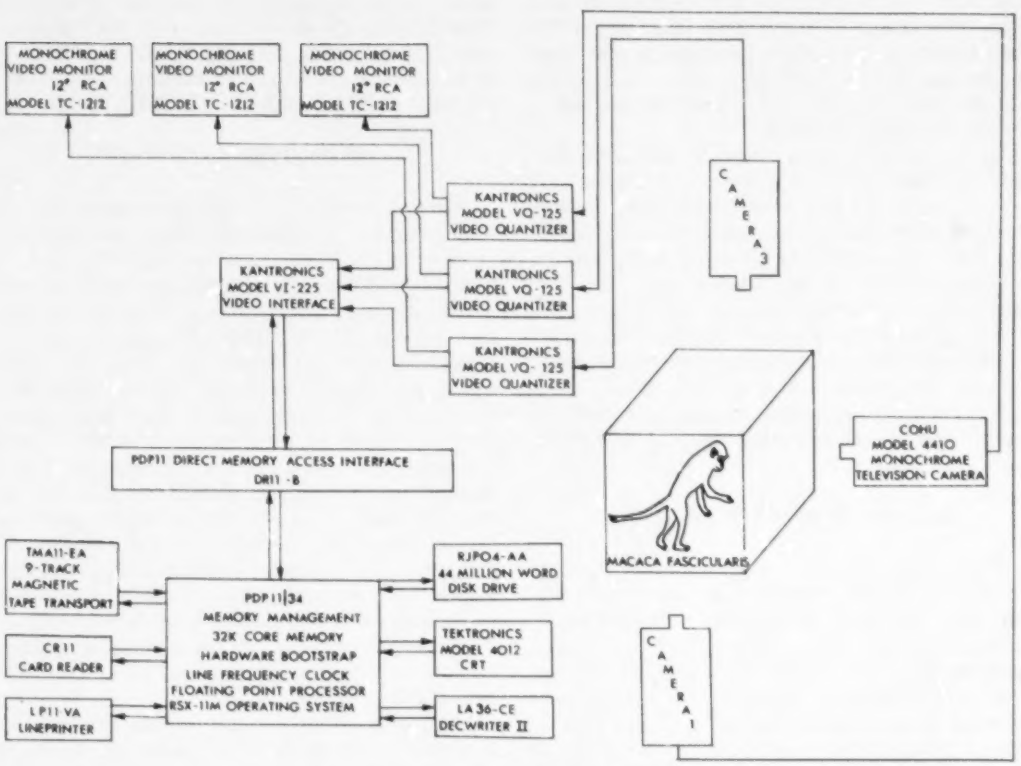
MONOCHROME VIDEO MONITOR 12" RCA MODEL TC-1212
MONOCHROME VIDEO MONITOR 12" RCA MODEL TC-1212
MONOCHROME VIDEO MONITOR 12" RCA MODEL TC-1212
KANTRONICS MODEL VQ-125 VIDEO QUANTIZER
KANTRONICS MODEL VI-225 VIDEO INTERFACE
KANTRONICS MODEL VQ-125 VIDEO QUANTIZER
KANTRONICS MODEL VQ-125 VIDEO QUANTIZER
CAMERA 3
COHU MODEL 4410 MONOCHROME TELEVISION CAMERA
MACACA FASCICULARIS
PDP11 DIRECT MEMORY ACCESS INTERFACE DR11-B
TMA11-EA 9-TRACK MAGNETIC TAPE TRANSPORT
RJP04-AA 44 MILLION WORD DISK DRIVE
PDP11/34 MEMORY MANAGEMENT 32K CORE MEMORY HARDWARE BOOTSTRAP LINE FREQUENCY CLOCK FLOATING POINT PROCESSOR RSX-11M OPERATING SYSTEM
CR11 CARD READER
TEKTRONICS MODEL 4012 CRT
LP11 VA LINEPRINTER
LA36-CE DECWRITER II
CAMERA 1

Figure 1. PROBE computer and video system hardware.

a specified voltage (greytone) range selected by the operator (see Figure 2). The video interface receives binary data from three video discriminators, converts them into a series of 16-bit data words, and transfers these data via a direct memory-access device to computer memory.

The video system is capable of processing images at four levels of resolution: 64 by 64, 128 by 128, 256 by 256, and 512 by 512 pixels (a pixel is a picture element). The degree of image resolution is set by a toggle switch in the video interface. This resolution setting directly determines the amount of detail present in the video image. Examination of very fine detail or small objects could require a 512 by 512 setting, whereas the outline of a larger object might only requirc a 64 by 64 pixel resolution.

Three critical variables interact in the processing of video data with PROBE: (1) image resolution, (2) image-sampling rate, and (3) data processing time. Increasing the resolution of the system by a factor of 2 increases the amount of data that must be processed by a factor of 4. For example, changing the resolution from 64 by 64 to 128 by 128 pixels increases the data generated for one picture by one camera from 4,096 to 16,384 bits. The number of times an image is taken is a function of the observation situation and the limitations of the system hardware. In the experiments completed to date (observing a fast-moving primate), a relatively rapid sampling rate (2 samples/sec) was necessary to avoid losing data. Although it would be ideal to process video images on the basis of necessary sampling rate and resolution, the data handling capacity of the computer system must also be considered. Video transfer rates are limited by the memory cycle speed on the host computer. The PDP-11/34 operates effectively at a data transfer rate of one 16-bit word every 1.22 microsec without a loss of data. In addition to transferring the image data to computer memory, time must be allowed

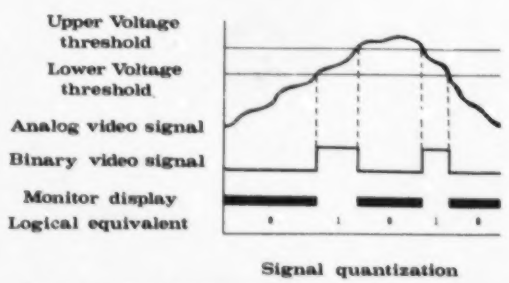

Figure 2. The relationship among the video signal, the voltage thresholds, and the equivalent bit structure.

for the movement of data to a storage device prior to taking another image. These processes, therefore, are the limiting factors in the hardware controlling how much image data can be taken per unit of time. To date, this system has been set to a .5-sec sampling rate with a resolution of 128 by 128 pixels.

There are three separate cameras, and when data coming from only one of these are referenced, the word "view" is used; however, when data from all three cameras are referenced simultaneously, the word "frame" is used. Thus a frame of data is the sum of the information in three views. In the present usage, the word "image" is less precisely defined. In general, image is used to mean the form of the information before it has been processed into binary data by the discriminator circuits and the interface. In this sense, "image" occasionally is used to mean that fraction of a view that contains the shape of the animal, that is, the nonzero vectors.

OBSERVATION ENVIRONMENT

Defining the conditions of observation was critical to the design of PROBE. Providing an environment that would allow the maximum diversity of spontaneously occurring behaviors while keeping the observation situation as simple as possible was the initial criterion.

The basic environment consists of a 3 x 3 m room constructed of plywood paneling with sound-attenuating insulation packed between the walls. The interior of the room is painted white and illuminated by six two-unit fluorescent lights. The observation cage is 88 cm square at the top and bottom and is a 68 x 68 cm rectangle on each side. The cage is constructed of 3.8-cm angle aluminum and .48-cm aluminum bars spaced at 4.4-cm intervals. Access to the cage is through a sliding panel of bars located in the center of one side wall. Two of the video cameras are placed approximately 1.6 m from the center of the sides of the cage; the third camera is located 2.4 m above and centered on the ceiling of the cage (Figure 3). The cage and its platform are painted white, thereby providing a background with a maximum degree of contrast with a dark-colored primate.

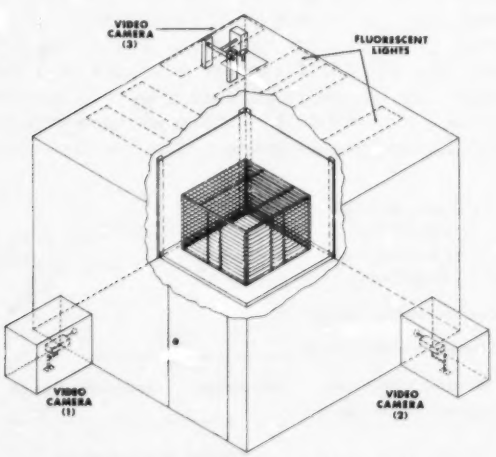

Figure 3. A line drawing of the observation environment.

BEHAVIORAL TAXONOMY

Work on a behavioral taxonomy began with a team of four observers independently viewing videotapes of the primates in the observation environment. After many hours of individual observation, the team as a group developed verbal descriptions of the motor activities of the primates. Using these descriptions, operational definitions were developed for what the observers agreed were discrete behavioral activities. Then observation sessions of 13 min duration[1] each were carried out on a number of the primates using PROBE. A program was written so that each of these frames could then be displayed on the graphics oscilloscope. The observation team examined each observation session and classified the behavioral acts displayed by the primates. A few acts originally recognized with continuously running videotapes were no longer observed when the data-taking procedure was limited to one frame every .5 sec. These acts were then eliminated from consideration.

This classification process resulted in the following list of behavioral acts.

Primary positions: *Crouch*—At least one hind foot on the floor, stomach in close proximity to the floor. *Sit*—Posterior in close proximity with horizontal surface, not crouching. *Quadrupedal Stand*—Stationary, three or four feet on the floor. *Bipedal Stand*—Stationary, one or both hind feet on the floor, major body axis approximately vertical. *Hand Stand*—Both forefeet on cage floor, hind feet not on the floor. *Cling*—Any nonperching, nonhanging position with two or more feet on a vertical cage surface, vertical body position, feet not on the floor. *Stretch*—A "cling" with the major body axis in the horizontal plane. *Perch*—Posterior on the wall or in the cage corner, not on the floor; body in the cage corner, Camera Views 1 and 2 showing a profile with legs jutting out. Monkey facing away from the cage corner. *Hang*—Only two feet on the cage ceiling, hind feet not on the floor. *Quadrupedal Hang*—Three or four feet on the cage ceiling.

Movement acts: *Quadrupedal Walk*—Horizontal movement in lower half of the cage, feet on the floor. *Inverted Quadrupedal Walk*—Quadrupedal hang with movement. *Bipedal Walk*—One or two feet on the floor, horizontal movement with head in the upper half of the cage. *Climb*—One or more feet on a vertical surface, movement in a vertical direction, vertical body position. *Inverted Climb*—A "climb" with an inverted body position. *Scale*—Any "cling" or "stretch" or "hang" with movement across a vertical cage surface. *Scoot*—A "sit" to "sit" with intervening horizontal movement.

Secondary acts: *Groom*—Hand or mouth in contact

with another body part, occurs in a "sit," "crouch," "quadrupedal stand," "bipedal stand," "cling," "perch," and "quadrupedal hang." *Reach*–Body weight supported by hind feet. Occurs with (1) "sit" or "crouch," either or both front appendages extended, elbow angle greater than or equal to 120 deg, trunk-arm angle greater than or equal to 45 deg; (2) "perch," either or both front appendages on ceiling; (3) "quadrupedal stand" or "bipedal stand," either or both front appendages on a vertical surface or ceiling. *Body Turn*–Body rotation about its vertical axis with movement of the feet; occurs with any "stand," "sit," "crouch," or "hang." *Body Shift*–A greater than or equal to 20-deg change in orientation from the vertical axis, occurs with a "sit," "bipedal stand," "cling," "perch," and "hang."

With the definition of mathematical properties associated with each act, the acts are defined in a logically exclusive way in the software. Our experience suggested that motor activity could be more easily determined if a monkey was initally identified as being in one of the 10 primary body positions. Each primary position could then be linked with one of four secondary activities, if present, to define an individual behavioral act. Selected primary positions were also used to determine if movement occurred from one observation (video frame) to the next. The assignment of a primary body position, secondary activity, or movement produced a taxonomy of 40 individual motor behaviors. Examples of the primary body positions are presented in Figure 4. These operationally defined behaviors are the acts that the PROBE system was programmed to recognize.

THE PATTERN RECOGNITION SYSTEM

The complete program system can logically be divided into the following phases: data acquisition, pattern analysis, and pattern classification. Data acquisition is the process of converting the data from the physical source (video image) into a form acceptable to the computer for further processing. Included in this phase is the preprocessing done by the discriminator circuits and any real-time computation performed on the data to simplify and reduce the volume of data to be stored on disk or magnetic tape. All of this must be accomplished during the observation session. Pattern analysis includes all computation used to identify the invariant characteristics of the original data and the associated pattern classes that might exist. Finally, pattern classification includes the computation needed to decide to which act each observation of the primate is to be assigned.

During the development of the pattern recognition system, these three processes must be carried out in a mixed order requiring some degree of iteration and interaction. For example, reliability measures of the pattern classification may well suggest a critical change in one or more parameters used in the pattern analysis process. Many such iterations may be required before the pattern recognition problem is satisfactorily solved.

All computational work for this project had to be accomplished using the minicomputer. This constraint required a major effort to make every program efficient in terms of computing time. Often several experiments are run concurrently, and all can be in the data acquisition phase. This results in entire working days being devoted to 13-min observation sessions, during which each observation session produces almost 5 million words of data to be processed and stored. The system must, in addition, provide the computation for pattern analysis and classification, as well as later statistical analysis for toxic effects, without developing a serious backlog of uncompleted computation. With this computing load, efficiency in computing time is mandatory.

The data acquisition program is important but straightforward. Each .5 sec the computer sends a signal to the video interface. This signal forces the three cameras to simultaneously begin a new view, thus insuring synchronization of the video data. The 3,072 words of data that correspond to the three video views are temporarily stored in arrays indexed by view number. The data are then transformed from a bit-by-bit format into a vector of the nonzero values of the bits for each master line of the video scan. For subsequent computations, the data used are these nonzero line vectors.

The pattern classification phase, described below, is essentially a mathematical model of the acts of the monkey. This model considers the original image as being represented by a feature vector of N numerical values. This vector identifies a point in the N-dimensional classification space, in which each element is a measure of some property of the original image. The pattern analysis software calculates the elements of this vector. The design problem for the pattern analysis program is to mathematically describe or represent the human definitions of behavioral acts. If an exact representation of the human definition is not possible, then features must be constructed that convey sufficient information to imply the same classification.

Mathematical Techniques

The data from one camera view can mathematically be treated as a matrix of (128 by 128) resolution cells or pixels obtained by sampling and digitizing the original image. Each element of this matrix will have a value of 0 or 1, corresponding to the original image element. This makes it convenient to regard a digitized view as a real-valued nonnegative function of two variables, f(x,y). The principal task is to define a set of mathematical properties of the function f(x,y) that either duplicate the human definition of a given class or convey sufficient information to imply the same classification. The selection of a set of properties for use in such an effort

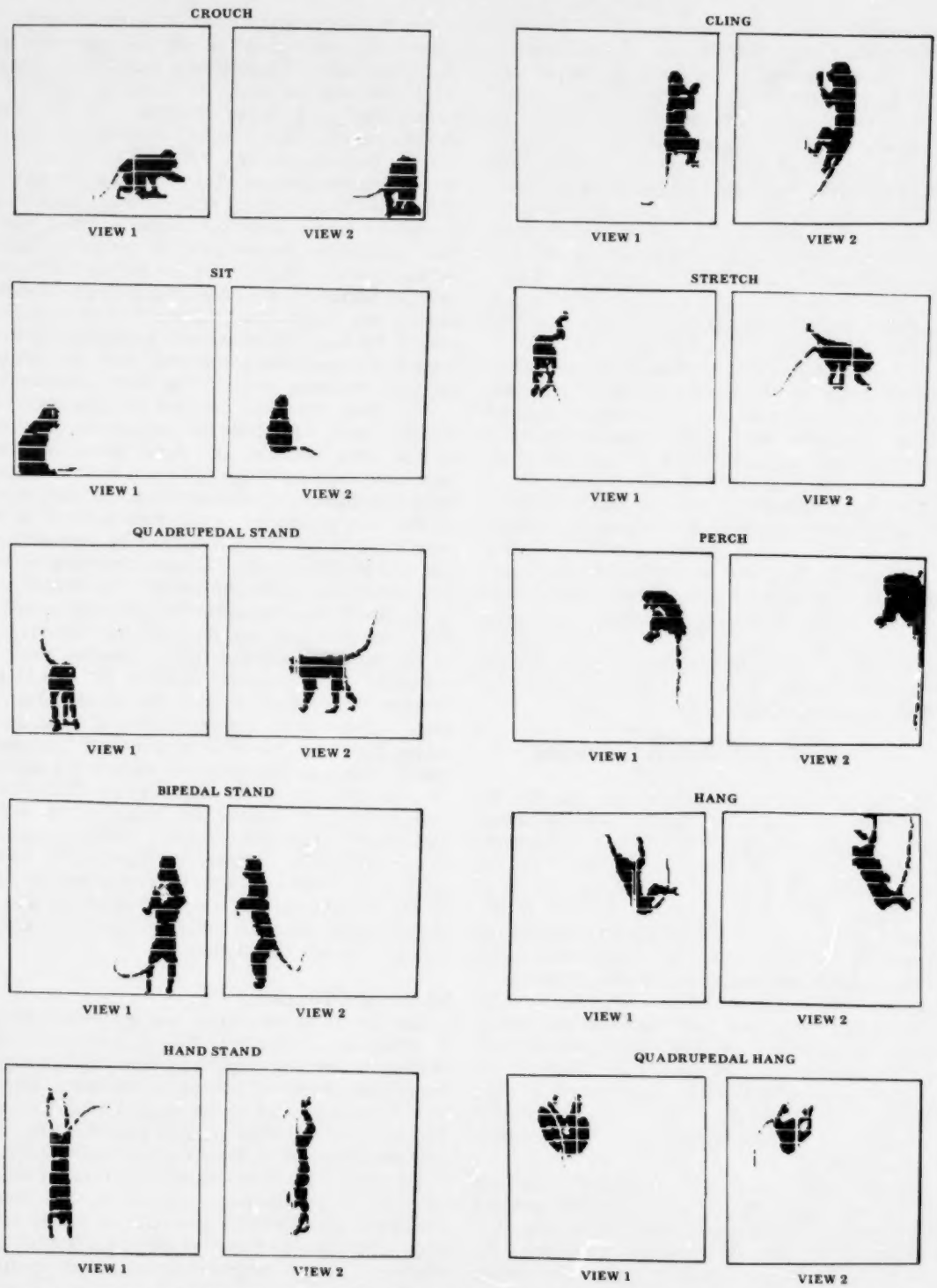

Figure 4. Examples, as seen from Video Cameras 1 and 2, of the 10 primary behavioral acts as reproduced by PROBE on a storage display scope.

is mainly an intuitive choice on the part of system designers. The technique of property selection is based on the method of lower order moments defined as:

$$M_{\alpha\beta} = \int\int x^{\alpha} y^{\beta} f(x,y)\,dx\,dy. \quad (1)$$

Moments have been used previously for ship and aircraft identification (Dudani, Breeding, & McGee, 1977; Smith & Wright, 1971). In the special case in which f is a discrete function and the area of each cell of f is defined as 1, Equation 1 is:

$$M_{\alpha\beta} = \Sigma_x \Sigma_y x^{\alpha} y^{\beta} f(x,y). \quad (2)$$

The lower order moments are then:

$$M_{00} = \Sigma_x \Sigma_y f(x,y), \quad (3)$$

$$M_{10} = \Sigma_x \Sigma_y x f(x,y), \quad (4)$$

$$M_{01} = \Sigma_x \Sigma_y y f(x,y). \quad (5)$$

The total number of dark pixels (area) in the picture is M_{00}. For example, if f is of Order 3 by 3, then M_{10} is:

$$M_{10} = \Sigma_x x\,[f(x,1) + f(x,2) + f(x,3)]. \quad (6)$$

A simple way of visualizing this is to move a vertical slit horizontally across the picture. At each x_i, form the sum of all dark pixels in the slit. This sum can then be considered as a weight, W_i, associated with the x_i. These weights are the projections of the picture f(x,y) onto the x-axis, and the moments are simply expressed as:

$$M_{00} = \Sigma_i W_i, \quad (7)$$

$$M_{10} = \Sigma_i W_i x_i, \quad (8)$$

$$M_{01} = \Sigma_j W_j y_j. \quad (9)$$

This process of collapsing onto the x- and y-axes is believed to play a major role in the discrimination of visual patterns by the octopus (Young, 1964, pp. 127-137).

A useful property for a general location of the body mass of the primate is expressed by the center-of-gravity coordinates:

$$\bar{x} = \frac{\Sigma_i W_i x_i}{\Sigma_i W_i} \quad (10)$$

and

$$\bar{y} = \frac{\Sigma_j W_j y_j}{\Sigma_j W_j}. \quad (11)$$

Another useful property is based on the sum of squares of weighted deviations from the mean ($\bar{x}$), given by:

$$SS = \sum_{i=1}^{n} (x_i - \bar{x})^2 W_i. \quad (12)$$

This provides a measure of dispersion. If the x_i values in the projection tend to be close, SS will be small; however, if they scatter over a wide range, SS will be large. The average deviation (standard deviation of the variable x_i) is:

$$\sigma = \{[\sum_{i=1}^{n} (W_i) - 1]^{-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 W_i\}^{1/2}. \quad (13)$$

Since for any view it is useful to project on two separate axes, x and y, the notation σx or σy will be used to automatically indicate which average deviation is being used. Some simple shape features can be obtained by comparison between the deviation of the x-axis projection (σx) and the deviation of the y-axis projection (σy). If σx is larger than σy, then the shape is more elongated in the horizontal direction. This is generally true when the primate is in a "crouch" position. The opposite is true for "bipedal stand" and erect "sit" positions. In these cases, the posture is more elongated in the vertical direction, resulting in a σy that is larger than σx.

Theoretical Discussion

The technique used in the pattern classification process is a combination of two methods that are often referred to as a "decision-theoretic method" and as "discriminant techniques" (Sebestyen, 1962, p. 61). The general approach of each method and the manner in which the methods are merged for solving the present problem are described below.

If a feature vector $V = (V_1, V_2, \ldots V_n)$ is used as a mathematical model to represent an original image, then each observation corresponds to a point in this N-dimensional space. In any space, independent of the number of dimensions, a vector of the type used in this discussion goes from the origin of the space to a particular point. Therefore, there exists a logical relationship between the vector and a unique point in the space. Two vectors can be compared by mathematical tests upon the vectors themselves, for example, by examining the vector product of the two vectors or by examining the spatial separation of the two points that correspond to the two vectors. It is expected that the points corresponding to different events, each belonging to the same class, will cluster in the N-dimensional space such that their separation will on the average be small. The separation of each of these points from the region occupied by a second class is expected to be larger if this technique is to be useful in distinguishing to which of these two classes an act belongs. This classification approach has been referred to as a decision-theoretic method (Sebestyen, 1962). As the number of possible

classes becomes large, the dimensions of the required feature space also increase.

Methods that do not rely on similarity measurements but exploit the dissimilarities that exist between different pattern classes have been referred to as "discriminant techniques." These methods employ techniques that discriminate between one class and all remaining classes. The use of discriminant techniques for classification has definite advantages in the present application. The decision process can be accomplished by a series of binary class-separation rules, since the separation is always of one class from all remaining classes to be considered. For example, the position of the center of mass of the animal can be used to separate acts occurring on the floor from those occurring on the walls or the top of the cage. Different feature vectors can be formed and used at each step of the decision process. For each feature vector used, the major property differences between the class under consideration and all remaining classes would be exploited.

In merging these approaches, the acts are divided initially into broad categories corresponding to cage position, for example, those acts that occur on the floor or on the cage walls or on the cage top. This decision level utilizes discriminant techniques. For classifications within a broad category, appropriately defined feature vectors are used in a decision-theoretic technique.

Examination of Subregions

A human description of a particular behavioral act focuses on specific regions of the image. The observation of three or more legs on the floor, an arm reaching out, and so on, all contribute to the final classification of the act. The center of gravity $(\bar{x},\bar{y})$ is used as the focal point for locating specific body regions. For example, the head will be above $(\bar{x},\bar{y})$ when the primate is sitting, and the legs will be below $(\bar{x},\bar{y})$ for a quadrupedal stand position. The question then becomes how far below, above, left of, or right of $(\bar{x},\bar{y})$ are the particular local regions of interest?

The standard deviation, σ, is not only a measure of the dispersion of a given distribution, but it can also serve as an indicator of how far around $(\bar{x},\bar{y})$ to search for particular local regions of interest. A normal distribution with a mean μ and deviation σ has 68% of its area located in the region $(\mu - \sigma)$ to $(\mu + \sigma)$ A uniform distribution has approximately 58% of its area within this region. Based on these assumptions, the majority of the body mass of the primate should be contained in the region $(\bar{y} - 2\sigma y)$ to $(\bar{y} + 2\sigma y)$, where $\bar{y}$ and σy are the mean and deviation of the shape distribution projected onto the y-axis. The image can now be considered as consisting of four specific regions (see Figure 5). Two regions are above $\bar{y}$ and two are below $\bar{y}$. When these various regions of the image are used in the discussion of classification, the notation of Figure 5 will be used, in which Region 1 means the highest part of the image and Region 4 means the lowest part of the animal image.

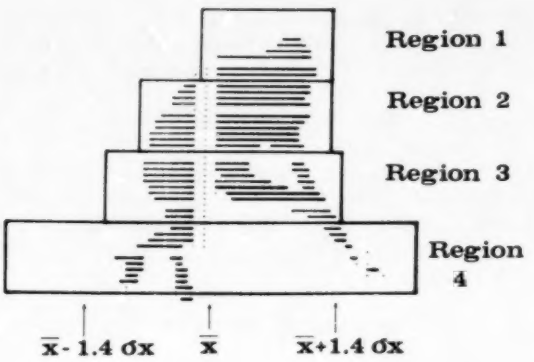

Figure 5. An example of how PROBE divides a view into four regions located symmetrically about the center of mass.

A feature vector for the primary floor acts is formed using the four body regions for each horizontal view. The properties associated with each element are as follows: V_1, legs present in the lower body region; V_2, the overall erectness of the shape of the animal; V_3, the (horizontal) x-axis spread of the shape in Region 3 of the view as compared with the same spread in Region 4; V_4, the compactness of Region 4; and V_5, the percent of the total shape that is contained in Region 4.

Linear combinations of these elements for both Cameras 1 and 2 are generally used in the classification decisions. The notation V_i^k is used for this feature vector, where i is the element number described above and k is the view number. As an example, the logic to classify an act among the possible major floor acts is described. The general techniques for deciding other major acts (that is, those that occur on the wall or on the ceiling) and among the modifiers (reach, turn, shift, groom) are quite similar. In some situations, it is useful to divide the image into segments or regions and to examine some or all of the regions in detail, as has been shown in Figure 5.

Classification of Floor Acts

Acts classified as quadrupedal stands or crouches are first separated by applying the criteria indicated in Figure 6 for the feature space of V_1 and V_2. The relationships between the physical posture of the animal and V_1 and V_2 are easily expressed. As $(\Sigma_k V_1^k)$ increases, this corresponds to an increase in the indication of the presence of legs in the floor region. A value of 3 or greater for $(\Sigma_k V_1^k)$ is a positive indication of the detection of three or four legs. As the animal assumes a more erect posture, the value of $(\Sigma_k V_2^k)$ will decrease. Therefore, a relatively large value for V_2^k indicates a very nonerect position. If the act is to be classified as a crouch, an additional requirement is imposed; namely, V_4 for both Cameras 1 and 2 must be less than .1.

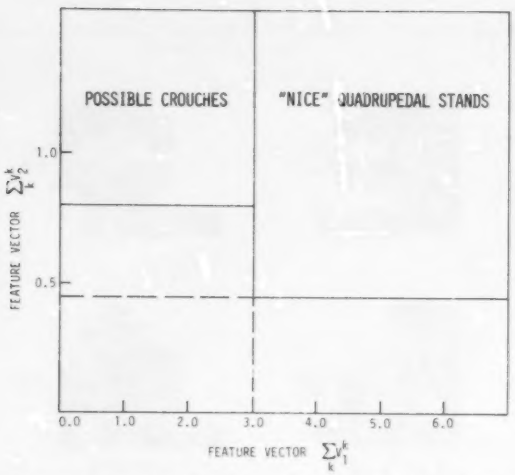

Figure 6. First separation of quadrupedal stands from crouches.

In considering the remaining acts, the initial decision is whether the data of a frame should be considered "erect" (sits and bipedal stands) or "nonerect" (quadrupedal stands, sits, and bipedal stands). The erect acts are determined by requiring that V_2 for both horizontal views be less than .35. The separation of erect sits from erect bipedal stands is accomplished by using the combination of $\Sigma_k(V_1{}^k + V_3{}^k)$. V_1 and V_3 are in the same units, since each is a ratio of σx computations. When V_1 is large, it reflects a considerable separation between the legs of a bipedal stand. An erect sit generally has a larger distribution along the x-axis of the body mass in Region 4 of the view than does the same distribution considered only for Region 3 of the view. This causes the $(\Sigma_k V_3{}^k)$ to be small. Figure 7 shows the criteria used for the separation of these erect acts. The overlapped area in which the variable $\Sigma_k(V_1{}^k + V_3{}^k)$ has a value between 4.5 and 5.0 is caused by sits that do not involve both forearms being relatively near the torso but during which at least one arm is somewhat extended and is touching the floor as a means of support. The acts in this area are separated by requiring that a bipedal stand must have V_3 greater than 1.0 in both horizontal views.

The nonerect floor acts are separated initially by deciding whether the data correspond to a bipedal stand. Bipedal stands are required to satisfy each of the following three tests: (1) $(\Sigma_k V_5{}^k)$ is less than .40; (2) in both horizontal views, V_2 must be less than .45; and (3) a particular horizontal view is chosen that has the smaller value of V_2 (the shape is most erect). In this view, V_1 must be greater than 1.0 or V_3 must be greater than 1.2.

The remaining nonerect acts of sit and quadrupedal stand are separated by requiring that $(\Sigma_k V_2{}^k) > .8$ for sits. Any remaining floor acts will then be classified as a quadrupedal stand.

Explanation

The detailed mathematical methods of treating the data in order to produce these vectors have been discussed earlier and are essentially methods using the lowest order moments of the distribution. In addition to the moment, the standard deviation, sigma, of the distribution is also used. In general, σ measures the variation of the distribution when projected onto an axis. For many decisions, the view is broken up into regions around the central point. Moments and standard deviations of the various regions in each view are also used in the decision process. This technique allows the majority of the decisions to be made in terms of ratios of moments or ratios of standard deviations. Depending upon the particular decision, the ratio used may involve data from different segments of one view, it may be a ratio of a variable on the x-axis to the same variable for the y-axis, or it could be a ratio of a variable as calculated for each of the horizontal views. The use of ratios has a major advantage for the present application because it allows the decision making aspects of the program to be independent of the absolute size of the animal being studied. This use of ratios in the decision process also allows the same basic program approach whether we are dealing with large monkeys or small monkeys. This is important in the present applications because the behavior of infant monkeys is also being studied.

It is possible to describe the various vectors in the N-dimensional space used to classify an act with a verbal description that is understandable to an individual who has no background in pattern recognition. For example, the variable V_1 measures the presence of legs in the lower body region. It does this by reflecting the presence of a dark area followed by a white area, followed by another dark area when the lower body region is scanned visually. This can be seen by comparing the diagrams shown in Figures 8 and 9. In each of these diagrams, the original shape is shown under "a." If a is projected onto the x-axis, it produces the figure shown in Figures 8b and 9b. If a is projected onto the y-axis, it produces the figure shown in Figures 8c and 9c. If c is

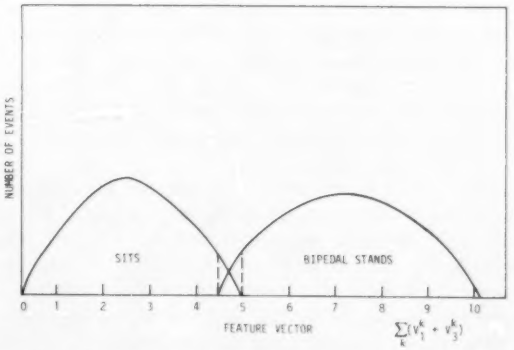

Figure 7. Separation of sits and bipedal stands for "erect" acts.

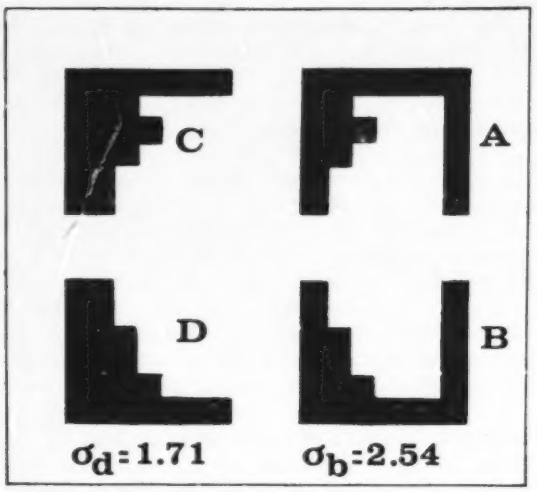

Figure 8. Representation of a nonsolid shape.

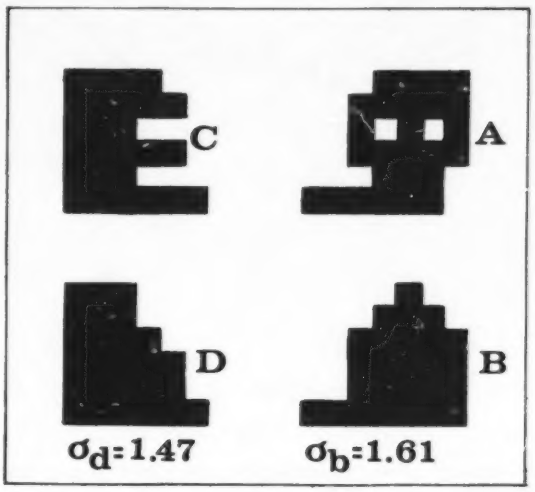

Figure 9. Representation of a solid shape.

in turn projected onto the x-axis, the result is as shown in Figures 8d and 9d.

Figure 8a has a structure that would be equivalent to two legs showing in the lower region of a body; Figure 9a is the same percentage of dark area but distributed differently. The standard deviations of Figures 8d and 9d are not appreciably different. This is because the process of projecting onto the y-axis first and then onto the x-axis loses the information contained in the original diagram (that is, either Figure 8a or 9a). However, the standard deviations of Figures 8b and 9b are significantly different. The comparison of the moments of the distributions and the standard deviations clearly indicate that Figure 8a amounts to dark areas that are separated by a large white area and that Figure 9a does not correspond to this type of a distribution. The process of going from Figure 8a to Figure 8c and then to Figure 8d produces what we refer to as a "transformed" distribution. If the standard deviations of the x-axis of Figure 8b are divided by the standard deviation for Figure 8d, the ratio is about 1.5. For Figure 9, this ratio would be approximately 1. The value near 2 represents the presence of large light areas between two major dark areas in Figure 8; the value near 1 for Figure 9 represents the absence of this feature. It is in this sense that this ratio can be used to detect the presence of legs in the lower body region.

The variable V_2 is used as a measure of the erectness of the major body features in the original frame. This is essentially the ratio $\sigma x/\sigma y$ for the entire image. The variable V_3 is the ratio $(\sigma x^T)_3/(\sigma x^T)_4$, where the superscript T means that the distribution has been transformed as discussed in the explanation of Figure 8d. Therefore, V_3 measures the ratio of the width of the shape in Region 3 of the image to the width of Region 4. The variable V_4 is a measure of the compactness of the lower region of the body mass.

The major variables that are calculated from the original vector data and that will act to classify the behavioral act accordingly to a particular frame of data are presented in Table 1. In Table 2 are listed the variables and combinations of variables used to decide categories of acts (for example, floor acts as opposed to nonfloor acts) or to decide between particular groups of acts (for example, sit vs. bipedal stand). No attempt has been made to reproduce the complete logic of the decision process or the detailed values of the variables that have been used to describe a particular act. The complicated nature of the decision process is indicated by the fact that four different tests are shown in Table 2 associated with the separation of the act sit from the act bipedal stand.

These techniques form the basis for the structure of the N-dimensional feature space. With these techniques, the major acts are correctly identified with an accuracy comparable to that of an experienced human observer. The Kappa coefficient for two experienced observers, each classifying 24 observation sessions, was approximately .85; the Kappa value for each observer with the PROBE classification for these sessions was .68. If the analysis is restricted to major acts, that is, modifiers such as groom, turn, and shift are ignored, the comparison between the computer results and human observation improves.

RECORD STRUCTURE AND DATA ANALYSIS

Many applications of pattern recognition have no problems associated with long-term storage of new data or of computed results. The situation is quite different in this application. An experiment may be as short as 1 month or as long as 2 years. When the data collection stage of an experiment is complete, all classification records must be available for analysis. In addition, it is

Table 1
Major Classification Variables

Variable	Description
$\bar{x}(j)$	Center of gravity of the x-coordinate for View j
$\bar{y}(j)$	Center of gravity of the y-coordinate for View j
$\sigma x(j)$	Standard deviation of $\bar{x}$ for View j
$\sigma y(j)$	Standard deviation of y for View j
$\bar{x}_i(j)$	$\bar{x}$ for the ith region for View j*
$\bar{y}_i(j)$	$\bar{y}$ for the ith region for View j*
$\sigma x_i(j)$	σx for the ith region for View j*
$\sigma y_i(j)$	σy for the ith region for View j
$\bar{x}^T(j)$	Center of gravity for the "transposed" distribution of View j as shown in Figure 8c or 9c
$\sigma x^T(j)$	Standard deviation associated with $\bar{x}^T$ above for View j
$p(j)$	Percentage of light in the lower body region for View j*
$M_i(j)$	Percentage of the body mass in the ith region for View j*
$h(j)$	Height of the body as shown in View j
$\Delta\bar{x}(j)$	Change in the x-center of gravity for View j from one frame to the next
$\hat{\sigma}x_i(j)$	"Normalized" value of $\sigma x_i(j)$, defined as $\hat{\sigma}x_i(j) = \sigma x_i(j)/\sigma x(j)$

**For an example in which the view is divided into different regions and each region is examined individually, see Figure 5.*

Table 2
Variables Used to Decide Classifications

Variable	Description of Test	Act
$y + 2\sigma y$	Near floor test	Floor/Nonfloor
$\sigma x(1)/h(1) + \sigma x(2)/h(2)$	Horizontal orientation	Crouch or Quad Stand
$p(j)$	Stomach near floor	Crouch
$\sigma x(1)/\sigma x^T(1) + \sigma x(2)/\sigma x^T(2)$	Clear presence of legs	Quad Stand
$\sigma x(j)/h(j)$	Vertical orientation	Erect/Nonerect
$\bar{x}_3(j)/x_4(j)$	Leg detection	Sit/Bipedal Stand
$\sigma x(j)/\sigma x^T(j)$	Leg detection	Sit/Bipedal Stand
$\sigma x(j)/\sigma x^T(j) + \bar{x}_3(j)/\bar{x}_4(j)$	Leg detection	Sit/Bipedal Stand
$M_4(1) + M_4(2)$	Leg detection	Sit/Bipedal Stand
$[\Delta\bar{x}^2(1) + \Delta\bar{x}^2(2)]^{1/2}$	Movement between adjacent frames	Movement
Compare $\sigma x(1)$, $\sigma x(2)$	Turn for floor acts	Turn
$\sigma x(j)/\sigma x^T(j)$	Arm separated from lower body for sits	Reach
$\tan^{-1}\{\sigma y(j)/[\bar{x}_i(j) - \bar{x}_1(j)]\}$	Tilt in the body orientation	Shift
Compare $\hat{\sigma}x_i(j)$, $\hat{\sigma}x_k(j)$	Relationship of normalized deviations for different regions	Reach
$\sigma x(1)/h(1) + \sigma x(2)/h(2)$	Horizontal orientation for Views 1 and 2	Hang
$\sigma x_1(j)/\sigma x_1^T(j)$	Arms in upper body region	Hang
Compare $\bar{x}_1(j)$, $\bar{x}_2(j)$	Position of head	Perch
$[\sigma x_3(j) + \sigma x_4(j)]/\sigma y(j)$	Feet protruding away from body	Perch
$\sigma y(j)/\sigma y^T(j)$	Legs near cage wall	Perch
$\lvert\bar{x}_1(j) - \bar{x}_2(j)\rvert > \lvert\sigma x_1(j)\rvert$	Reach detection	Perch

necessary to maintain, in a reasonably compact form, all the raw data taken with the PROBE system. This allows for the reanalysis of data either to detect previously unexpected effects or to improve sensitivity to drug effects if major improvements have been made in the classification program. Reclassification of data after improvements in the pattern recognition and classification programs has proven extremely useful. For example, during the analysis of data on monkeys exposed to d-amphetamine, unexpected indications of stereotypic cage behavior were found, and these methods of data storage and retrieval allowed a detailed analysis of this effect to be performed.

Long-term storage is maintained on magnetic tape. The record format for this corresponds to 1,200 (16-bit) words per frame (three views) of initial data. Of this, 1,000 words are devoted to storage of the original vector information for the frame. The remaining 200 words contain identification information and all calculated secondary quantities for that frame. If reclassification is required, this is the storage level and medium used.

For analysis of an experiment, a shorter record format is used that corresponds to 32 (16-bit) words per frame of data. All identification information is retained, as well as the classification, center of gravity, and standard deviations for each view. Normally, this short record is maintained on disk with a tape backup. Most data analysis is done utilizing only this short data record.

PRESENT STATUS AND FUTURE DEVELOPMENT

The PROBE system is complete and fully operational. With this system, data for experiments on the effect of

d-amphetamine and chlorpromazine on spontaneous behavioral activity have been taken. The results of these experiments will be published elsewhere (Hopper, Kernan, Higby, Cunningham, Lloyd, & Reiter, Note 1). At present, five major experiments are underway utilizing the PROBE system, including experiments on prenatal lead exposure and two experiments each with methylmercury hydroxide and thallium sulfate.

If we were to redesign the PROBE hardware, we would implement a number of changes. The bit-packing arrangement would be changed so that a computer word corresponded to data from one view only, thereby eliminating the necessity for later unpacking and shuffling of the data for storage in a view-by-view arrangement. In addition, the success of the vector approach, as contrasted to an earlier bit-by-bit approach to the video data, strongly suggests that a vector-generating preprocessor should be incorporated into the system. The inclusion of these simple design changes in the system would improve the efficiency of the system by decreasing the demand upon the central processing unit of the host minicomputer.

Some additional refinements in the classification programs should be implemented to improve the reliability of the classifications. In addition, some investigation utilizing human observers should be undertaken to test the completeness of the present taxonomy. In the work to date, there are indications that the PROBE system is sensitive to the existence of acts that were not distinguished as separate behaviors in the early human observation work. For example, the PROBE results indicate that there may be three different acts, all of which the human observer classifies as "sit," but which are separable on the basis of postural position. These preliminary indications should be examined in detail. There is some indication in our work that these posturally separate "sits" are affected differently by drug exposure. If these indications are confirmed, the sensitivity of this system will be both demonstrated and improved.

It is obvious that a great deal of work remains to be done if the full potential of this or similar systems is to be realized. The possible applications of PROBE-type system to other areas of behavioral analysis seem limitless. PROBE can be used to study the behavioral effects of toxic agents, neurological diseases (e.g., Parkinson's disease), social behaviors, or the correlation of behavioral and neurophysiological phenomena.

REFERENCE NOTE

1. Hopper, D. L., Kernan, W. J., Higby, W. J., Cunningham, W., Lloyd, W. E., & Reiter, L. *PROBE: Pattern recognition of behavioral events*. Manuscript submitted for publication, 1980.

REFERENCES

Dudani, S. A., Breeding, K. J., & McGee, R. B. Aircraft identification by moment invariants. *IEEE Transactions of Computers*, 1977, **C-26**, 39-46.

Fu, K. S. *Sequential methods in pattern recognition and machine learning*. New York: Academic Press, 1968.

Fukunaga, K. *Introduction to statistical pattern recognition*. New York: Academic Press, 1972.

Pavlidis, T. *Structural pattern recognition*. New York: Springer, 1977.

Sebestyen, G. S. *Decision-making process in pattern recognition*. New York: Macmillan, 1962.

Smith, F. W., & Wright, M. H. Automatic ship photo interpretation by the method of moments. *IEEE Transactions of Computers*, 1971, **C-20**, 1089-1095.

Young, J. F. *A model of the brain*. London: Oxford University Press, 1964.

NOTE

1. In theory, any length of observation period is possible. We have fixed upon an arbitrary choice of a 13-min observation session, and this has been used in all of the work to date. With this time, the data from eight observation periods fit on one 2,400-ft magnetic tape.

(Received for publication May 19, 1980;
accepted June 27, 1980.)